Some analytical foundations of multidimensional scaling for ordinal data

Feger, Hubert

SOME ANALYTICAL FOUNDATIONS OF MULTIDIMENSIONAL SCALING FOR ORDINAL DATA

HUBERT FEGER

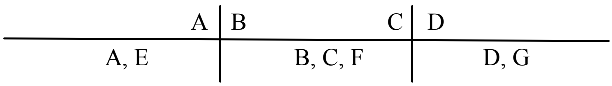
Abstract: *Ingwer Borg has contributed intensively and successfully to MDS, in theory and applications (e.g., Borg 1981a,b; Borg & Lingoes 1987; Roskam, Lingoes & Borg 1977). This paper offers some notes on the foundations of MDS, based on ranks of proximities. Two approaches are sketched, one working with contingencies of distance ranks, represented by boundaries in a dimensional space. The other approach uses the generalized betweenness relation, leading to configurations of object points. Details of the procedures and examples for both approaches are given for the one- and two-dimensional case. A procedure to find an optimal solution in a given dimensionality for data with random error is illustrated. The role of facet theory for theory testing by MDS is emphasized. Using the concepts of this paper will allow a fine- grained evaluation of a MDS solution for ordinal data.*

Space as structured by boundaries between points

Placing a point on a line divides the line into two parts. The point functions as a boundary between these parts. We will assume that a point on a line has two sides. Placing two points, A and B, on the same line, a third point may be found with equal distance to both, called by Coombs (1964) – in the context of unfolding theory – the 'working midpoint'. We write this point as A|B for the interval AB between the points A and B. Generalising to **k** > 1 and, with **k** the number of dimensions, one may call this separating boundary the mid-perpendicular hyperplane.

By **N** points $\binom{N}{2}$ boundaries are created. Each boundary divides all points into two sets. E.g. for A|B, some points are closer to A than to B, being located on the A-side of A|B, written C – A|B. The other points are closer to B and thus on the B-side of A|B. In general, two boundaries divide all points into four sets, each point is an element in two sets. On the other hand, two boundaries on a line define three intervals, two open and one closed. How are the two sets assigned to the three intervals? Figure 1 illustrates one example.

Figure 1 Four sets of points in three intervals on a line



On the A-side of A|B we find {A, E}, on the B-side {B, C, D, F, G}. On the C-side of C|D one sees {A, B, C, E, F}, and on the D-side {D, G}. Such a representation is possible only if one logically possible set of points does not exist. Here this empty set is defined by points on the A-side of A|B and simultaneously on the D-side of C|D. No interval is provided for such a location on the line in Figure 1. The distribution of the points relative to two boundaries may be described by a contingency table. An equivalent to Figure 1 is the fourfold classification in Table 1. The rows refer to the relative distance between points. A point is either closer to A or to B of A|B, written AB if it is closer to A. At the same time, a point is either closer to C or to D. The cells contain the points under consideration. Of course, A must be placed in the *AB*-row because in a metric space it is closer to itself than to B. If the order is A – B – C – D, the assignments in Table 1 result.

Table 1 Contingency table for the distribution of points

	CD	*DC*
AB	A	
BA	B, C	D

One cell, called the *zero cell*, is empty. In Table 1, the zero cell is specified as AB, DC. This is a necessary condition for representing all points on one line, i.e., to give a one-dimensional representation or seriation: For all pairs of pairs of N points, at least one cell in the contingency table is empty.

With more than two points, at least three boundaries and three pairs of boundaries exist. A consistent order between these boundaries must be found with respect to sequence and orientation to each other. The sequence may be written either from 'left to right' or from 'right to left'. The orientation of two boundaries is defined by referring to the fact that the other boundary is either on the one or the other side. E.g. for A|B – A|C, the other boundary of A|B is on its B-side; for A|C, it is on its A-side. The sides of the boundary defining the closed interval are oriented towards each other, or ‘inwards’. The other sides are oriented ‘outwards’. The rule how to represent the boundaries on a line is: Let a zero cell be defined as **A**B, **D**C. The first element in each defining pair is placed outside. For Table 1 we derive: A|B – C|D.

To test the consistency of seriation and orientation of A|B, A|C, and B|C for A – B – C, with AB < BC, we first write the three contingency tables (see Table 2).

Table 2 Contingency tables to test consistency of boundary orientation

[1]	*AC*	*CA*
AB	A	
BA	B	C

[2]	*BC*	*CB*
AB	A	
BA	B	C

[3]	*BC*	*CB*
AC	A, B	
CA		C

The last table shows two zero cells which lead to two solutions: A|C – B|C or C|A – C|B. Taking the first partial solution of [3], the order of the partial solutions from the three tables is consistent (see Table 3). One of the several axiomatizations of the betweenness relation may be used to test consistency routinely.

Table 3 Consistency demonstration

from [1]	A|B	-	A|C		
from [2]	A|B	-	-	-	B|C
from [3]			A|C	-	B|C
	A|B	-	A|C	-	B|C

This analysis corresponds to Feature Pattern Analysis (FPA, Feger 1994, Feger & Brehm 2001) and will now be applied to an example from Borg & Groenen (1997: 4).

Order by boundaries, illustrated by an example

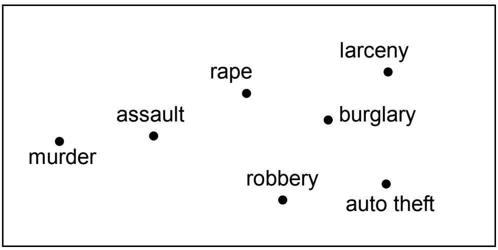
Table 4 reports from Borg & Groenen (1997: 4) the Pearson correlations between rates of different crimes over 50 US states (Wilkinson 1990). The ranks (with ties) of these coefficients are given below the main diagonal.

Table 4 Correlations between rates of different crimes (above main diagonal) / ranks of coefficients (below main diagonal)

		A	B	C	D	E	F	G
murder	**A**	-	.52	.34	.81	.28	.06	.11
rape	**B**	12.5	-	.55	.70	.68	.60	.44
robbery	**C**	16	10.5	-	.56	.62	.44	.62
assault	**D**	1	3.5	9	-	.52	.32	.33
burglary	**E**	19	5	6.5	12.5	-	.80	.70
larceny	**F**	21	8	14.5	18	2	-	.55
auto theft	**G**	20	14.5	6.5	17	3.5	10.5	-

Borg and Groenen provide in their Figure 1.1 a two-dimensional MDS representation of the correlations, see Table 4. We reproduce this solution in Figure 2.

Figure 2 Borg & Groenen solution for the correlations in Table 4



One may gain the impression that a part of the solution is one-dimensional: A – D – B – F. To test this assumption, we calculate all 15 contingency tables for these four elements. We find that all contain at least one zero cell. The order of the 6 boundaries is consistent, as graphed in Figure 3.

Figure 3 The one-dimensional order of four crimes

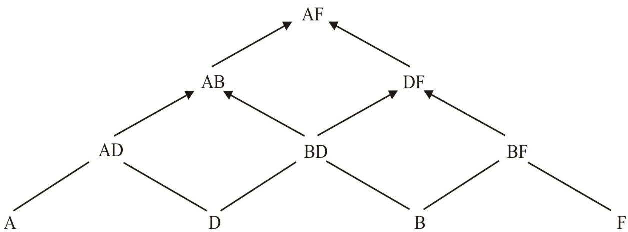
boundaries A|D D|B A|B A|F D|F B|F

points **A** | **D** | | | **B** | | **F**

Figure 3 locates the points within their intervals. B has to be located between A|B and B|F. But referring to the information in Table 4, its position can be determined more precisely. The figure is not drawn to proportion, e.g. D – D|B ≠ B – B|D.

Quantitative information. Knowing the orientation of two boundaries towards each other one can derive a comparison of the relative length of two distances. The distance between the points defining the inner sides of the boundaries is smaller than the distance between the points defining the outer sides, e.g. for A|B – C|D, we derive that BC < AD. To test the consistency of this quantitative information, one may use the *Pyramid Criterion*. In the example with the sequence A – D – B – F as the qualitative solution, the adjacent intervals are AD, BD, and BF. Combining two adjacent intervals to a longer new one, containing both, we derive AD + BD = AB and BD + BF = DF. Finally, the extreme points define the longest distance. Figure 4 represents this as a graph.

Figure 4 Graph of the Pyramid Criterion for the A – D – B – F example



In Figure 4, AD → AB means AD < AB. Transitive closure is implied. Furthermore, if AD < BF then AB < DF because both, AB and DF, include BD and without it correspond to AD or BF. In the example, the distances satisfy the pyramid criterion with AB < DF and AD < BF.

We are now ready to construct the ‘quantitative solution’ by solving a set of equations and inequalities derivable from the qualitative solution in Figure 3 and the ranks in Table 4:

$$AD + BD = AB \tag{1}$$
$$BD + BF = DF \tag{2}$$
$$AD + BD + BF = AF \tag{3}$$
$$\text{as side constraints: } AD < BD < \ldots < AF$$

One solution is AD = 4, BD = 6, and BF = 8. The quantitative solution with these distances is given in Figure 5. The quantitative information in this figure is exact but the figure is not drawn to proportion. The upper part of Figure 5 shows the positions of the points and the interpoint distances. The lower part provides the location and orientation of the boundaries and the distances between the boundaries and between the points.

Figure 5 A quantitative solution for A – D – B – F

A 4 D 6 B 8 F

2 A|D 2 |1 A|B 2 B|D 2 A|F 1 | 1 D|F 3 B|F 4 |

Even if, as in the crimes example, an acceptable **k** = 1 solution for all elements does not exist, one-dimensional *partial structures* may be found to support the interpretation. One of several such structures for the crimes data is shown in Figure 6.

Figure 6 A partial one-dimensional structure within a *k* = 2 solution

A	D	A	B	A	G	B	F
		A	C	B	E	C	G
		A	E	D	C	D	F
		A	F	D	E		

A	***D***	***B***	***C***	***E, F, G***

The first contrast, A|D, stresses that A (murder) is unique in comparison to all others. Closest to murder is assault. These two are different from all others, separated by four polarities: A to B (rape), to C (robbery), to E (burglary), and to F (larceny). One should not forget that this order is not based on similarity judgements but on common causes for the occurrence frequencies in the US states. Again, four contrasts separate A, D, and B from the others, all referring to violence against human beings versus violations of property rights.

Two-dimensional generalisation: trivariate contingencies

Again, from Figure 2 we hypothesize that the points for D = assault, B = rape, C = robbery, and E = burglary are arranged in two dimensions. If this is true, then either no zero cell exists in at least one contingency table, or the partial solutions derived from these tables are not consistent. We find that several tables require a **k** > 1 solution. Some of these tables are reported in Table 5.

Table 5 Contingency tables requiring *k* > 1

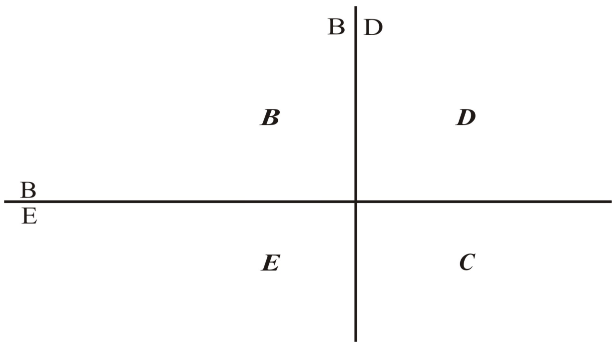
[1]	*BE*	*EB*
BD	B	E
DB	D	C

[2]	*CD*	*DC*
BD	E	B
DB	C	D

[3]	*DE*	*ED*
BD	B	E
DB	D	C

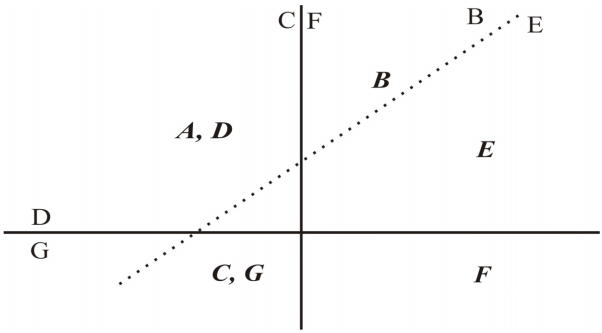
In **k** = 2, boundaries are separation *lines*. Two lines must intersect to provide four quasi-quadrants to locate each of the four points separately (see Figure 7, representing subtable [1] of Table 5).

Figure 7 Two separation lines locating four points



Referring to Table 5, it can be seen that some boundaries partition the points in the same way, e.g. B|C = D|C = D|E. We now turn to all points simultaneously. For N = 7 there exist 21 pairs of points. To every pair corresponds a boundary line, partitioning the set of all points into two sets. The sets are defined by their relative distances to the two points that define the boundary. Figure 8 shows the intersection of C|F and D|G. Then B|E is added to gain more information about the location of the seven points.

Figure 8 Intersection of boundary lines to locate all points



If such a map can be drawn without inconsistency, the combination of the three boundaries involved has passed the consistency test. Should such a test be failed, a way out would be to increase the number of dimensions. The information in Figure 8 can be written as a 2 x 2 x 2 contingency table (see Table 6). As before, BE means: The distance of the column point(s) to B is shorter than to E, etc.

Table 6 Contingency table for three boundaries

			points
BF	CF	DG	**A, D**
		GD	zero cell <1>
	FC	DG	**B**
		GD	zero cell <2>
FB	CF	DG	zero cell <3>
		GD	**C, G**
	FC	DG	**E**
		GD	**F**

This contingency table shows eight rows which are the cells of this table. In the example, three zero cells exist. They are represented as in the one-dimensional case with the first element defining a pair oriented outwards. The map in Figure 5 corresponds to zero cell <2>. The existence of more than one zero cell allows more than one representation or demonstrates a lack of uniqueness for the location of (too few?) points, given these boundaries. In general, the uniqueness increases rapidly with **N**.

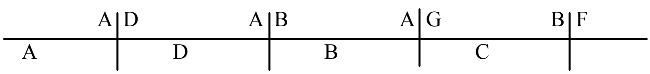
Without going into constructional details, we mention that not every intersection of three boundary lines leads to seven (isotonic) regions. If only three points define all three boundaries, as in A|B, A|C, and B|C, a closed region results corresponding to points with distances AB < BC, BC < AC, but AC < AB, an intransitive order. Unless one wants to represent intransitive choices, this region can be condensed to a point. This results in a star-like intersection of the three boundaries, corresponding to the intersection of mid-perpendiculars of a triangle in Euclidean space.

In the example depicted in Table 4, with N = 7, there exist 21 boundary lines A|B … F|G. Table 7 lists the partitions into two sets of elements that each line achieves.

Table 7 Partitions by boundary lines, crimes example

	boundary	*partition*	*comment*
(1)	A|B	A, D |B, C, E, F, G	same as A|C, A|E, A|F
(2)	A|D	A |B, C, D, E, F, G	
(3)	A|G	A, B, D |C, E, F, G	same as B|E, C|D, D|E
(4)	B|C	A, B, D, E, F |C, G	
(5)	B|D	A, C, D |B, E, F, G	same as C|E
(6)	B|F	A, B, C, D |E, F, G	same as C|G, D|F
(7)	B|G	A, B, D, F |C, E, G	
(8)	C|F	A, C, D, G |B, E, F	
(9)	D|G	A, B, D, E |C, F, G	
(10)	E|F	A, B, C, D, E, G |F	
(11)	E|G	A, B, C, D, E, F |G	
(12)	F|G	B, D, E, F |A, C, G	

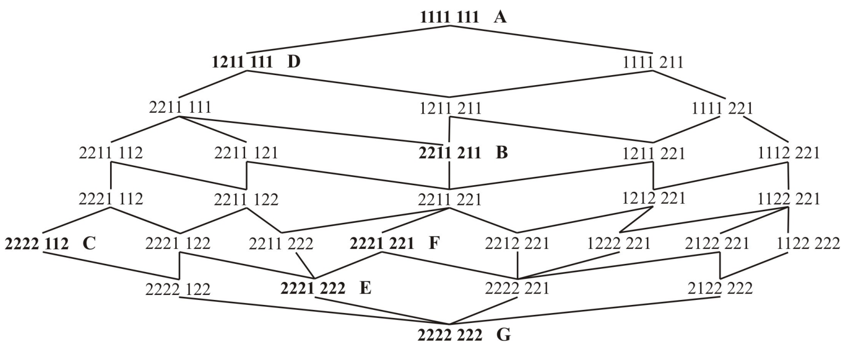
As Table 7 shows, only 12 out of 21 boundary lines provide a *unique* partition. For the *interpretation* of a solution, the fact that A|B and A|C lead to the same partitions provides the information that – as far as these data are concerned – the difference between murder and rape creates the same set structure among the elements as the difference between murder and robbery. Some lines may be drawn as pseudo-parallel, i.e., the data do not force these lines to intersect. In the example, several collections of such lines exist, hinting to one-dimensional substructures, as



The sets created by these partitions are *not* clusters as they are defined quite often in the literature, with the elements in a cluster being closer together and having longer distances to elements outside the cluster (Feger 2004).

A boundary line is not necessarily a mid-perpendicular bisector. Depending on the density distribution of the object points, a boundary line may separate the points in the same way as a mid-perpendicular bisector but the line is not in the same position as the bisector. From this viewpoint, one may *interpret a boundary line as a dichotomous item.* To perform an FPA, the crimes would be the data-generating cases. Taking only the 12 non-redundant boundary lines, this leads to a 12 x 7 matrix. One immediately sees the problem: too few cases. With a computer program, these data result in more than one hundred perfect two-dimensional solutions. Just for purposes of illustration, we report one solution with the first seven boundary lines in Figure 9. Note that this FPA-solution is compatible with the MDS-solution in Figure 2. The observed patterns (crimes) are printed bold. Could the other patterns in this solution – and they are the ones that differ between the solutions – be interpreted as other crimes?

Figure 9 An FPA solution with boundary lines as items



To the same 12 x 7 data matrix, other models can be fitted, e.g. HOMALS (see SPSS 12.0). The Eigenwerte of the first and second dimension are 0.408 and 0.307. The solution is very similar to the one in Figure 2, except for the position of D, e.g. with (ADBE).

Space as structured by distances between points

We now approach our problem from a different perspective (the discussion will compare the two approaches). Three distances are defined between three non-identical points. Without ties, one distance is the longest. If the points are collinear, the defining end points of the longest distance become the two extreme points, the third point is located *between* the other two. For each triple, this betweenness order can be derived. The properties of the one-dimensional betweenness relation (see, e.g. Fishburn 1985 for a review of axiomatizations) may be used to test consistency in the on-dimensional case.[1]

Let us briefly demonstrate the procedure for the crimes A, B, D, and F. The ranks of the proximities between these points are contained in Table 4. We derive

triple	*betweenness position*
ABD	A – D – B
ABF	A – B – F
ADF	A – D – F
BDF	D – B – F

leading to A – D – B – F. This is the same qualitative solution as in Figure 3, which was derived from the contingencies approach.

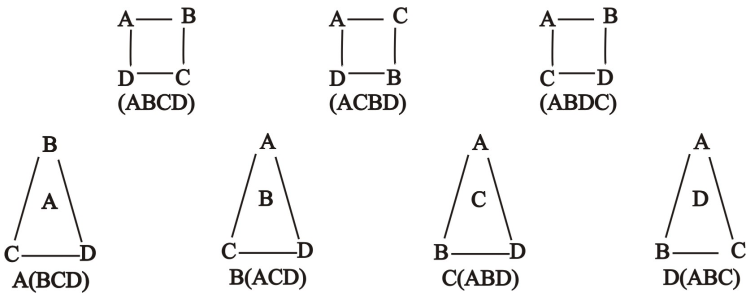
Two-dimensional generalization by using a multidimensional betweenness relation

We treat the multidimensional case by *generalizing the betweenness relation.* For **k** = 2, a point D is located between three other points A, B, and C if its position is on the same side of a line through A and B as point C, of a line through A and C as B, and on the same side of a line through B and C as A (see Figure 10).

On the plane, four points A … D may generate two topologically different configurations, either a (true) quadrilateral or a triangle with an inner point, i.e., one point between the other three points. Labeling the points, the seven configurations of Figure 10 result. This figure also reports the notational abbreviations of these configurations (see also Feger 2001).

1 written by Dipl. Math. Philip Metzner

Figure 10 All labeled configurations of four points in a plane



If the ranks of the six distances between these points are observed or derived, only some distance matrices are compatible with some configurations. Together with an extension to **k** = 3, Feger (1996) derived that

(1) any pair of two opposite sides of a quadrilateral is shorter than the sum of the two diagonals, and
(2) any two sides of a triangle are longer than those two inner lines originating from those end points which both triangle sides do not have in common.

For example in Figure 10, given (ABCD), we find (AB + CD) < (AC + BD) and (AD + BC) < (AC + BD). For A(BCD) one derives (BC + BD) > (AC + AD), (BC + CD) > (AB + AD), and (BD + CD) > (AB + AC). With these results, ranked distances can be used to identify the set of compatible configurations; their consistency can be tested by a constructive procedure to find the overall solution and will be presented below.

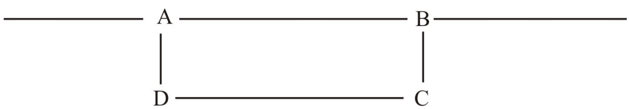
Testing for consistency (see Feger 2001). For the points A … E in Table 4 we list all compatible quadruple configurations in Table 8.

Table 8 Compatible configurations for A … E of Table 4

	quadruple	*configurations*		
		.1	.2	.3
1	ABCD	(ABCD)	D(ABC)	
2	ABCE	(ABEC)	B(ACE)	E(ABC)
3	ABDE	(ABED)	D(ABE)	
4	ACDE	(ADCE)	(ACED)	D(ACE)
5	BCDE	(BDCE)	B(CDE)	

For quadruple A, B, C, and D two compatible configurations are derived, to be numbered 1.1 and 1.2. The number of compatible configurations may differ from quadruple to quadruple. A solution consists of one configuration per quadruple. Thus, the number of combinations to be tested in this example is 2 x 3 x 2 x 3 x 2 = 72. While the sequence of this testing is irrelevant, we start with the combination of configurations 1.1, 2.1, 3.1, 4.1, and 5.1. Quadruple 1.1 corresponds to (ABCD). A line can be drawn through A and B extending this side of the quadrilateral. Then the remaining points C and D must be located on the same side – i.e., in the same half space – created by this line (see Figure 11).

Figure 11 The line through A and B creates two half spaces



The AB-line is also a part of 2.1 (ABEC). Here again, the points C and E have to be on the same side of this AB-line. Because C already is located on one side, C, D, and E must all be on this side, to be tested and confirmed by 3.1 (ABED). So the result of the first part of the consistency test is – [AB] C, D, E. In this example $\binom{N}{2} = \binom{5}{2} =$ 10 lines exist, each one to be used in the consistency test. For the line through A and C, we find

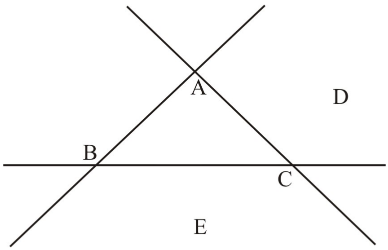
from	1.1	B [AC] D
	2.1	B, E [AC] –
	4.1	E [AC] D

and in the same way one derives

from	1.1	– [BC] A, D
	2.1	E [BC] A
	5.1	E [BC] D

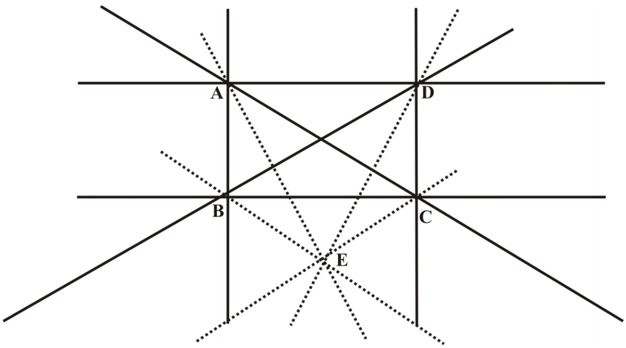
Together, and this reveals the constructive nature of this algorithm, we may draw Figure 12.

Figure 12 Integrating the results of the first three parts of the consistency test



The algorithm proceeds by using all lines and derives the solution in Figure 13, in which only the qualitative information counts.

Figure 13 One solution for A … E



The algorithm continues to find two more solutions (see Figure 14). All three qualitative solutions are in perfect agreement with the ranks of the distances. The first solution in Figure 14 is based on 1.2, 2.1, 3.1, 4.3, and 5.1, the second one on 1.2, 2.1, 3.2, 4.2, and 5.1. Being based on three identical configurations, their similarity is high; D is either located on the same side of line AE as C or as B.

Figure 14 Two more solutions for A … E

Comparing these solutions with Figure 13, (ABEC) is common to all three. The position of D is not well determined by this part of the data.

The analysis of the quadruple configurations can not only be used to construct a solution but allows a very *detailed study of the fit* between solution and data. Performing this analysis for the solution in Figure 2, two results are mentioned. (1) Identifying by inspection of Figure 2 the configurations implied, we note that most configurations in the solution are acceptable. Only one configuration violates the status of acceptability: We observe (ACGF), but (ACFG) would be acceptable, i.e., compatible with the ranks. For a solution with an almost perfect fit, it is not surprising that only one configuration does not correspond to the limitations provided by the data. (2) In Figure 2, point D is placed *on* the line between A and E. Our analysis indicates that D could be on either side. Thus, the solution in Figure 2 provides a 'compromise' of two equally acceptable qualitative solutions. The first contains D(ABE), (ADEC), D(AEF), and (ADEG,) while the second includes (ABED), D(ACE), (ADEF), and D(AEG). Thus, the quadruple analysis determines what is implied by the data – common over perhaps many equivalent solutions.

Facet theory. Schönemann & Borg (1983: 333) have emphasized that theoretical assumptions about the structure in the data should exist when performing MDS. One prominent possibility to relate the statistical analysis and the theoretical assumptions is facet theory. We illustrate one way to use quadruple configurations in **k** = 2 for theory testing with a well known example of eight intelligence tests (Guttman 1965; see also Borg 1992; Borg & Groenen 1997: 73). The correlations between the tests are given in the lower triangular matrix of Table 9.

Table 9 Correlations between intelligence tests

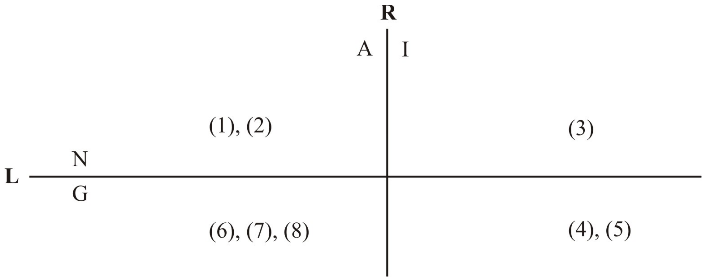
Tests	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
(1)	-							
(2)	.67	-						
(3)	.40	.50	-					
(4)	.19	.26	.52	-				
(5)	.12	.20	.39	.55	-			
(6)	.25	.28	.31	.49	.46	-		
(7)	.26	.26	.18	.25	.29	.42	-	
(8)	.39	.38	.24	.22	.14	.38	.40	-

The MDS solution for these data (Borg & Groenen, Figure 5.1) can be interpreted as being a circumplex. If this circular order exists in the data, every quadruple should exist in specific forms. For a circumplex without error with all points on a circle, e.g. the tests (1) … (4) should be represented by ((1)(2)(3)(4)). But the graph of the solution might suggest that test (6) deviates from a perfect simplicial structure. Testing the quadruple (4), (5), (6), and (7), one finds that the configuration compatible with the data is ((4)(5)(7)(6)) and not ((4)(5)(6)(7)), as demanded by a perfect circumplex. Furthermore, the quadruple analysis reveals that (4) or (5) or (6) may be between the remaining object points of this quadruple.

We conclude that the circumplex is not perfect and that the deviation of (6) is real and could deserve an explanation. We did not report the analysis of all 70 quadruple configurations. This is not necessary to reach the conclusions above, because if the acceptable configurations for A, B, C, and D do not contain the theoretically predicted one, no analysis including further object points E … N changes the list of acceptable forms for A … D. In this sense, the list of acceptable configurations for A … D is *independent* of any other list of quadruples from A … N. If a hypothesis is related to a subset of points, it suffices to analyze this subset. By the way, the solution reported in the literature is not totally supported by the data, e.g. ((4)(5)(6)(8)) does not exist. The quadruple analysis suggested here allows to pinpoint all deviations of a solution from the data.

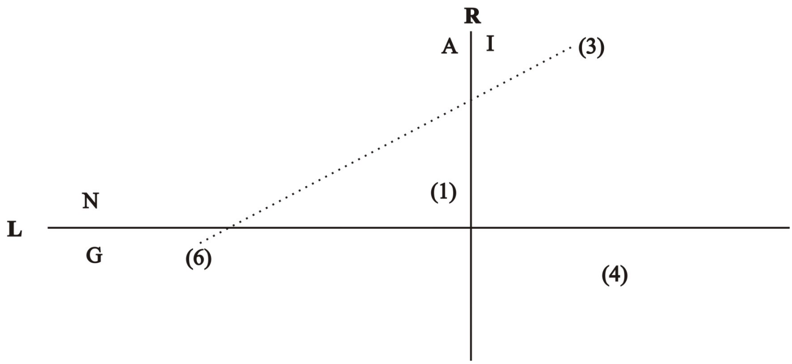
On the other hand, the eight tests are ordered by two facets, *language* **L** with N = numeric and G geometrical, and *requirement* **R** with A = application and I = inference. These two facets may be conceptualized and represented as boundary lines. Their intersection is provided in Figure 15.

Figure 15 Coding of the intelligence tests with respect to two facets



It is evident that all configurations maintaining the positions of the quadrants like ((1)(3)(4)(6)) or ((2)(3)(5)(8)) provide empirical support for the hypothesized facet structure. A configuration violating the clockwise (or anti-clockwise) order of the quadrants like ((1)(4)(3)(6)) would violate the assumed order. But the information that (1) is placed in the quadrant defined by N and A does not place (1) in an exact geometrical position. It is not against the information if (1)((3)(4)(6)) is derived from the data (see Figure 16). Thus, the violating cases are only quadrilaterals with diagonals against the quadrant order.

Figure 16 A betweenness position of test (1) compatible with the facet structure



Quantification. To demonstrate the quantification, given a qualitative solution, we select the crimes A … E and re-rank their distances in Table 10 (taken from Table 4).

Table 10 Re-ranked distances between crimes A … E

	A	B	C	D	E
A	-				
B	7.5	-			
C	9	6	-		
D	1	2	5	-	
E	10	3	4	7.5	-

We chose the first solution in Figure 14 for the demonstration, with D(ABC), (ABEC), D(ACE), and (BDCE) as the defining quadruples. For each quadruple configuration, we derive inequalities, e.g. for D(ABC) we derive (see paragraph 5 of this paper):

$$({}^{1}AD + {}^{2}BD) < ({}^{9}AC + {}^{6}BC),$$
$$({}^{1}AD + {}^{5}CD) < ({}^{6}BC + {}^{7.5}AB),$$
$$({}^{2}BD + {}^{5}CD) < ({}^{7.5}AB + {}^{9}AC).$$

Each of these three inequalities is trivial in the sense that for each rank in the smaller sum a larger rank exists – in a one to one assignment – in the larger sum. This observation helps to reduce the number of inequalities which increases very much with an increasing number of elements (crimes). These trivial inequalities are satisfied by the side constraints corresponding to the ranks as observed in Table 10: ${}^{1}AD < {}^{2}BD < \ldots < {}^{9}AC < {}^{10}AE$. Of all inequalities derived from the quadruple configurations, only

$$({}^{7.5}AB + {}^{7.5}DE) < ({}^{10}AE + {}^{2}BD)$$

is not trivial in the sense defined above and has to be taken into account when calculating the quantitative solution. Furthermore, to obtain metric distances, for every triple A, B, and C the triangular inequality has to be satisfied. For A, B, and C, with AC as the largest distance in Table 10, one requires AC ≤ AB + BC. Table 11 gives one of many possible sets of distances satisfying all constraints.

Table 11 Distances between A … E

	A	B	C	D	E
A	-				
B	24	-			
C	26	22	-		
D	13	14	20	-	
E	37	16	18	25	-

We did not prove that the conditions and constraints imposed on the distance estimates are sufficient. Therefore, we test the realizability of a two-dimensional representation of the distance estimates in Table 11 by applying MDS. Using PROCSCAL of SPSS 12.1 with a ratio transformation of the dissimilarities in Table 11, we obtained a normalized raw stress = 0.0011 and an explained variance of 0.99889. This corresponds as close as possible to the perfect fit to be expected in this analysis. The solution shows the correct forms of the quadruple configurations. The coordinates of the crimes A … E are provided in Table 12.

Table 12 Coordinates of A … E in the plane

	dimension	
	1	**2**
A	.845	-.029
B	-.142	.422
C	-.257	-.542
D	.330	.055
E	-.776	.095

Searching for an 'optimal' solution for data with error

For researchers applying MDS and aiming at a low dimensional solution, usually only solutions with some error or 'unexplained variance' exist, but not with a perfect fit between model and data. Sometimes, it is meaningful to assume that random error has distorted the size of the proximities and their ranks. Most conceptualizations of random error imply that small errors are more probable or frequent than large ones. For ranks, this means that the exchange of adjacent ranks (the first becomes the second, while the second becomes the first) should lead frequently to a dissolving of error, while the exchange of more distant ranks may not as often be necessary to find a solution.

We search for a solution with an optimal fit to the data. Of course, several definitions of 'optimal fit' can and have been given – implied in every optimization algorithm and choice of a loss function (see Borg & Groenen 1997). Our approach compares the distance ranks implied by a solution to those observed. We start by changing the observed distance ranks as little as possible. This we define in accordance with Kendall's τ as the number of exchanges or permutations between adjacent ranks. Kendall (1948) used the statistic **S** defined as the minimum number of exchanges of adjacent elements necessary to transform one rank order into another one. We illustrate our procedure with the ranks provided in Table 13.

Table 13 Distance ranks to demonstrate the approximation matrix

	A	**B**	**C**	**D**
A	-			
B	3	-		
C	1	5	-	
D	4	2	6	-

For these data, no perfect one-dimensional solution exists. With N objects and $\binom{N}{2}$ distances, there exist $\binom{N}{2}-1$ pairs of ranks of adjacent distances. For the example of Table 13 we derive this list of exchanges with **S** = 1:

	observed	*transformed*
[1]	^{1}AC, ^{2}BD,	^{2}AC, ^{1}BD.
[2]	^{2}BD, ^{3}AB,	^{3}BD, ^{2}AB.
[3]	^{3}AB, ^{4}AD,	^{4}AB, ^{3}AD.
[4]	^{4}AD, ^{5}BC,	^{5}AD, ^{4}BC.
[5]	^{5}BC, ^{6}CD,	^{6}BC, ^{5}CD.

We write the rank of a distance as an exponent to the left. ^{1}AC means that AC has rank 1, signifying the shortest distance. For each of the five exchanges with **S** = 1, we test whether a **k** = 1 solution exists. After exchange [1], we derive the ranks in Table 14.

Table 14 Distance ranks after exchange [1]

	A	**B**	**C**	**D**
A	-			
B	3	-		
C	2	5	-	
D	4	1	6	-

From the ranks in Table 14 and the triples B – A – C, A – B – D, C – A – D, and C – B – D, we construct the quantitative solution

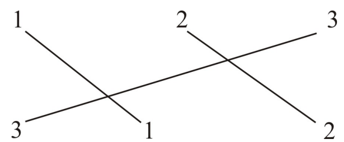
$$D \;\overset{1}{\text{————————}}\; B \;\overset{3}{\text{————————}}\; A \;\overset{2}{\text{————————}}\; C$$

This solution perfectly fits the ranks in Table 14, and approaches with **S** = 1 as close as possible the ranks in Table 13. Because there may exist more solutions with **S** = 1, i.e., the same optimal fit, we test the effects of all exchanges [1] … [5]. For exchange [2] we find the same solution as for [1], noting that slightly different rank matrices may lead to identical solutions. Exchanges [3] and [5] do not provide a solution, but [4] produces

$$D \;\overset{2}{\text{————————}}\; B \;\overset{3}{\text{————————}}\; A \;\overset{1}{\text{————————}}\; C$$

agreeing with the first solution qualitatively but not with respect to the quantification.

As far as the example is concerned, this would end the search for optimal solutions. In general, larger exchanges than **S** = 1 will be considered, first those with **S** = 2. These exchanges are generated by combining two **S** = 1 exchanges, e.g. [1] with [2], [1]with [3], etc. from the list above. Furthermore, **S** = 2 is created by a permutation of the type: from 1 – 2 – 3 to 3 – 1 – 2, illustrated as two intersections of the lines connecting the same ranks:



Discussion

Using the perspectives as emphasized by Schönemann & Borg (1983), we here describe and discuss the contributions of this paper with respect to the existence, uniqueness, and interpretation of nonmetric MDS solutions for ordinal data. Two formal approaches are outlined. The first one uses *contingencies* and, as the geometrical realization, *boundaries* between object points. The second approach applies the *generalized betweenness relation*. The geometrical correspondence is the seriation in a triple in **k** = 1 and the *quadruple configuration* in **k** =2. For both approaches, procedures and examples in the one- and two-dimensional case are provided.

Starting with **k** = 1 for the contingency approach, the *existence* of a solution depends on the existence of at least one zero cell per contingency table, and on the consistency of the zero cells. This leads to a qualitative solution in the sense of Coombs, with incomplete information on the (relative) size of the distances between object points. The existence of a quantitative solution requires the solvability of a system of equations and inequalities. To test this, the 'Pyramid Criterion' is introduced.

Using an equation system to find scale values meets some reservations (see Krantz et al. 1971, in particular chapter 9). Indeed, a larger system of empirically derived equations rarely is without a contradiction, rendering the whole system not solvable. General approximation procedures are available, e.g. Mathcad (2001) with the option *'minerr'*.

One may use the algorithm of Feature Pattern Analysis to find all qualitative and quantitative solutions. The scale level of the quantitative solutions approaches the interval scale with an increasing number of objects rather rapidly. Thus, the question of uniqueness is answered by listing all solutions compatible with the data and the acceptability criteria of the researcher.

For the interpretation of a solution, the concept of a boundary may be called upon. Such a boundary point or line or plane exists for every pair of object points and correspondents – substantially interpreted – to a *contrast* between the two objects defining this boundary. In the crimes example, some boundary lines express the difference between crimes with violence against human beings vs. violations of property rights.

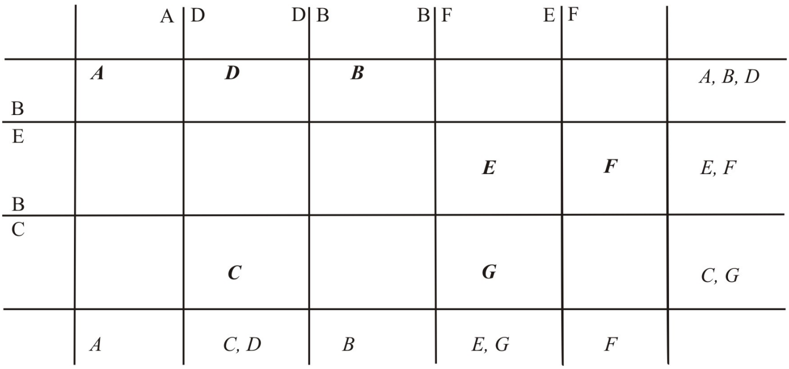
In the *two-dimensional case,* zero cells in trivariate contingency tables establish the necessary conditions for the existence of solutions. The structure implied by each zero cell must be compatible with the structure postulated by every other zero cell. Again, each pair of object points creates a boundary line, which now may intersect with other lines in the plane to create regions for the location of the object points. A boundary line may be interpreted as an item to be analyzed by procedures such as FPA or HOMALS. As in the one-dimensional case, all qualitative and quantitative solutions of a desired fit can be found.

The second approach, using proximity ranks in a one- or multidimensional *betweenness relation,* orders in **k** = 1 all triples of object points according to one of the various betweenness axiomatizations, and if that is possible simultaneously for all object points, a qualitative solution exists. The quantitative solution can be found in the same way as for the first approach. For the interpretation, each triple stipulates a comparison of the middle or central element with the two outer ones: The climate in the Netherlands (NL) differs from those of Sweden (S) and Portugal (P), but S and NL as well as P and NL have more or perhaps more important properties in common than S and P.

The two-dimensional case introduces the powerful concept of the *quadruple configuration.* Two topological variants exist in **k** = 2. The quadrilateral represents four object points, none of which is located between the others. The triangle with an inner point represents one object point between the other three. For more than four points, one may differentiate between the hull and the kernel of the constellation of points. Multiple imbedding of hulls and kernels allows for a detailed substantial interpretation. In Figure 2, for example, D (assault) and E (burglary) are contained in or surrounded by all other points, forming a hull for them. E also is contained in D, F, and G. With more points, the number of (relative) hulls with their kernels, forming overlays, tunnel, and other structures, provide a rich basis for interpretation, as the lines of the contingency approach do.

The positioning of co-ordinate axes is arbitrary, as in most MDS models. To support a dimensional substantial interpretation in **k** = 2, one of many possibilities is to use two sets of boundary lines. Each set intersects (only) with the other one. The sets should partition the point localizations in two different ways. These partitions should be as independent as the data allow. Independence may be defined, e.g. by Kendall's τ. Figure 17 reports one of many possibilities for the crimes data. The partitions in the last row and the last column (with many ties) correlate $\tau_b = 0.344$, which is not significant.

Figure 17 A two-dimensional representation of the crimes example



The two approaches lead to the same results, while using different ways. From the contingencies, zero cells are derived and geometrically realized as a set of boundaries (see Figure 1). From pairs of boundaries, one may then derive the positions of the object points in one of the regions. Quantitative information is provided by a zero cell as well. Using the betweenness relation on the distances, triple (in **k** = 1), quadruple (in **k** = 2) or quintuple configurations (**k** = 3; a tetrahedron with an inner point or two tetrahedra with a common triangular base) are the analytical units to be used when constructing a spatial representation. While the configurations immediately provide the localizations of the object points, the positions of the boundaries can be derived from them. Thus, in a different sequence, both approaches offer the same information on the location of points and of boundaries, starting both from ordinal data.

There exist, of course, other possibilities to provide a mathematical foundation for MDS. For metric MDS, several bases are already given (see Mathar 1997). Borg & Groenen (1977: 16) offer a “ruler-and-compass approach to ratio MDS”. The assumed properties of the data require different approaches. Other achievements are discussed in Schönemann & Borg (1983).

References

Borg, I. (1981a). *Anwendungsorientierte Multidimensionale Skalierung.* Heidelberg: Springer.

Borg, I. (Ed.) (1981b). *Multidimensional Data Representations: when and why.* Ann Arbor, MI.: Mathesis Press.

Borg, I. (1992). *Grundlagen und Ergebnisse der Facettentheorie.* Bern: Huber.

Borg, I., & Groenen, P. (1997). *Modern Multidimensional Scaling.* New York: Springer.

Borg, I., & Lingoes, J. (1987). *Multidimensional Similarity Structure Analysis.* Berlin: Springer.

Coombs, C. (1964). *A Theory of Data.* New York: Wiley.

Feger, H. (1994). *Structure Analysis of Co-occurrence Data.* Aachen: Shaker.

Feger, H. (1996). Positions of objects in space: Their definition and their effects on experience and behaviour. In W. Battmann & S. Dutke (Eds.), *Processes of the Molar Regulation of Behaviour* (pp. 55-72). Lengerich: Pabst Science Publishers.

Feger, H. (2001). Some applications of Feature Pattern Analysis in the social and behavioral sciences. *Psychologische Beiträge, 43*(2), 345-384.

Feger, H. (2004). Anmerkungen zur Clusteranalyse. In E. Witte (Ed.), *Methodologische, methodische und historische Entwicklungen in der Sozialpsychologie* (pp. 94-125). Lengerich: Pabst Science Publishers.

Feger, H., & Brehm, M. (Eds.) (2001). *New Developments in Feature Pattern Analysis. Psychologische Beiträge, 43(2).* Lengerich: Pabst Science Publishers.

Fishburn, P. C. (1985). *Interval Orders and Interval Graphs.* New York: Wiley.

Guttman, L. (1965). A faceted definition of intelligence. *Scripta Hierosolymitana, 14,* 166-181.

Kendall, M. G. (1948). *Rank Correlation Methods.* London: Griffin.

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of Measurement (Vol. 1).* New York: Academic Press.

Mathar, R. (1997). *Multidimensionale Skalierung. Mathematische Grundlagen und algorithmische Aspekte.* Stuttgart: Teubner.

Mathcad (2001). *Professional.* MathSoft Inc.

Roskam, E. E., Lingoes, J. C., & Borg, I. (Eds.) (1977). *Geometric Representations of Relational Data.* Ann Arbor, MI.: Mathesis Press.

Schönemann, P. H., & Borg, I. (1983). Grundlagen der mehrdimensionalen metrischen Skaliermethoden. In H. Feger & J. Bredenkamp (Eds.), *Messen und Testen. Enzyklopädie der Psychologie, Themenbereich B, Band 3* (pp. 257-345). Göttingen: Hogrefe.

Wilkinson, L. (1990). *SYSTAT: The system for statistics.* Evanston, IL.: SYSTAT Inc.